KU LEUVEN

Solving Probability Problems in Natural Language

Anton Dries, Angelika Kimmig, Jesse Davis, Vaishak Belle and Luc De Raedt

anton.dries@cs.kuleuven.be

https://dtai.cs.kuleuven.be/problog/natural_language

Our goal

Mike has a bag of marbles with 4 white, 8 blue, and 6 red marbles. He pulls out one marble from the bag and it is red. What is the probability that the second marble he pulls out of the bag is white?

The answer is 0.235941.

combination of mathematics
and natural language processing

Question answering

Challenges

1. Understand the question

2. Obtain background knowledge

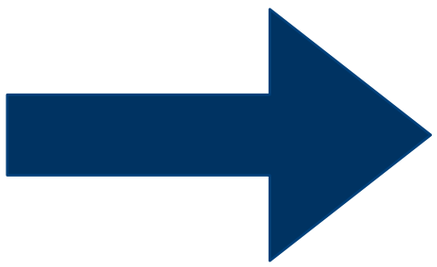
3. Solve the question

Question answering

Challenges

1. Understand the question

2. Obtain background knowledge



3. Solve the question

Inspiration

Aristo

Growing thicker fur in the winter helps some animals to

(A) hide from danger
(B) attract a mate
(C) find food
(D) keep warm

ARI

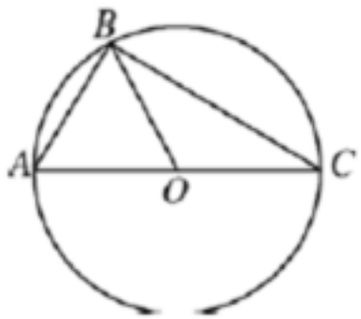
▸ (D) keep warm

OT

▸ (C) find food

some examples

In the figure above, triangle ABC is inscribed in the circle with center O and diameter AC. If AB=AO, what is the degree measure of angle ABO?



(A) 15°
(B) 30°
(C) 45°
(D) 60°
(E) 90°

Euclid / Geos

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

KU LEUVEN

Our approach

natural language

off-the-shelf NLP tools
+ rule-based system

specification language

solver

solution

Contributions

1. **Formal model** to represent probability questions

2. **Solver** to compute the solution of such a formal model

3. NLP component to **extract a formal model from text**

4. **Dataset** of over 2376 labeled questions

Example

Mike has a bag with 4 white,
8 blue, and 6 red marbles.

He takes one marble
from the bag and it is red.

What is the probability that
the second marble he takes
from the bag is white?

Example

Mike has a bag with 4 white, 8 blue, and 6 red marbles.

He takes one marble from the bag and it is red.

What is the probability that the second marble he takes from the bag is white?

setup

multiset bag

Values(color) = {white, blue, red}

#white(bag) = 4
#blue(bag) = 8
#red(bag) = 6

Example

Mike has a bag with 4 white,
8 blue, and 6 red marbles.

multiset bag

Values(color) = {white, blue, red}

#white(bag) = 4
#blue(bag) = 8
#red(bag) = 6

He takes one marble from the bag and it is red.

action + observation

first = take(bag)
#first = 1
rest(first) = bag \ first
observe #red(first) = #first

What is the probability that the second marble he takes from the bag is white?

Example

Mike has a bag with 4 white, 8 blue, and 6 red marbles.

multiset bag

Values(color) = {white, blue, red}

#white(bag) = 4
#blue(bag) = 8
#red(bag) = 6

He takes one marble from the bag and it is red.

first = take(bag)
#first = 1
rest(first) = bag \ first

What is the probability that the second marble he takes from the bag is white?

action + question

snd = take(rest(first))
rest(snd) = rest(first) \ snd
#snd = 1
probability #white(snd) = #snd

Example

Mike has a bag with 4 white, 8 blue, and 6 red marbles.

He takes one marble from the bag and it is red.

What is the probability that the second marble he takes from the bag is white?

multiset bag

Values(color) = {white, blue, red}

#white(bag) = 4
#blue(bag) = 8
#red(bag) = 6

first = take(bag)
#first = 1
rest(first) = bag \ first

observe #red(first) = #first

snd = take(rest(first))
rest(snd) = rest(first) \ snd
#snd = 1
probability #white(snd) = #snd

Examples

Other types of questions can be represented in the same formalism

A die is thrown 3 times.
Find the probability that the sum of the dots is at least 5.

A blood disease is found in 2 percent of the persons in a certain population. A new blood test will correctly identify 96 percent of the persons with the disease and 94 percent of the persons without the disease. What is the probability that a person who is called negative by the blood test actually does not have the disease?

Solver

ProbLog

```
% Probabilistic facts:
0.5::heads1.
0.6::heads2.

% Rules:
someHeads :- heads1.
someHeads :- heads2.

% Queries:
query(someHeads).
```

probabilistic logic programming language

“Prolog + probabilities”

easy integration background knowledge as logical rules

reasons over possible worlds

Solver

Direct encoding in ProbLog

```
:- use_module(library(lists)).

take(1, L, X) :-
    select_uniform(1, L, X, _).
take(I, L, X) :-
    I > 1,
    I2 is I - 1,
    select_uniform(I2, L, _, R),
    take(I2, R, X).

evidence(take(1, [w,w,w,w,b,b,b,b,b,b,b,b,r,r,r,r,r,r], r)).
query(take(2, [w,w,w,w,b,b,b,b,b,b,b,b,r,r,r,r,r,r], w)).
```

select_uniform(+ChoiceID, +List, -Element, -RestOfList)

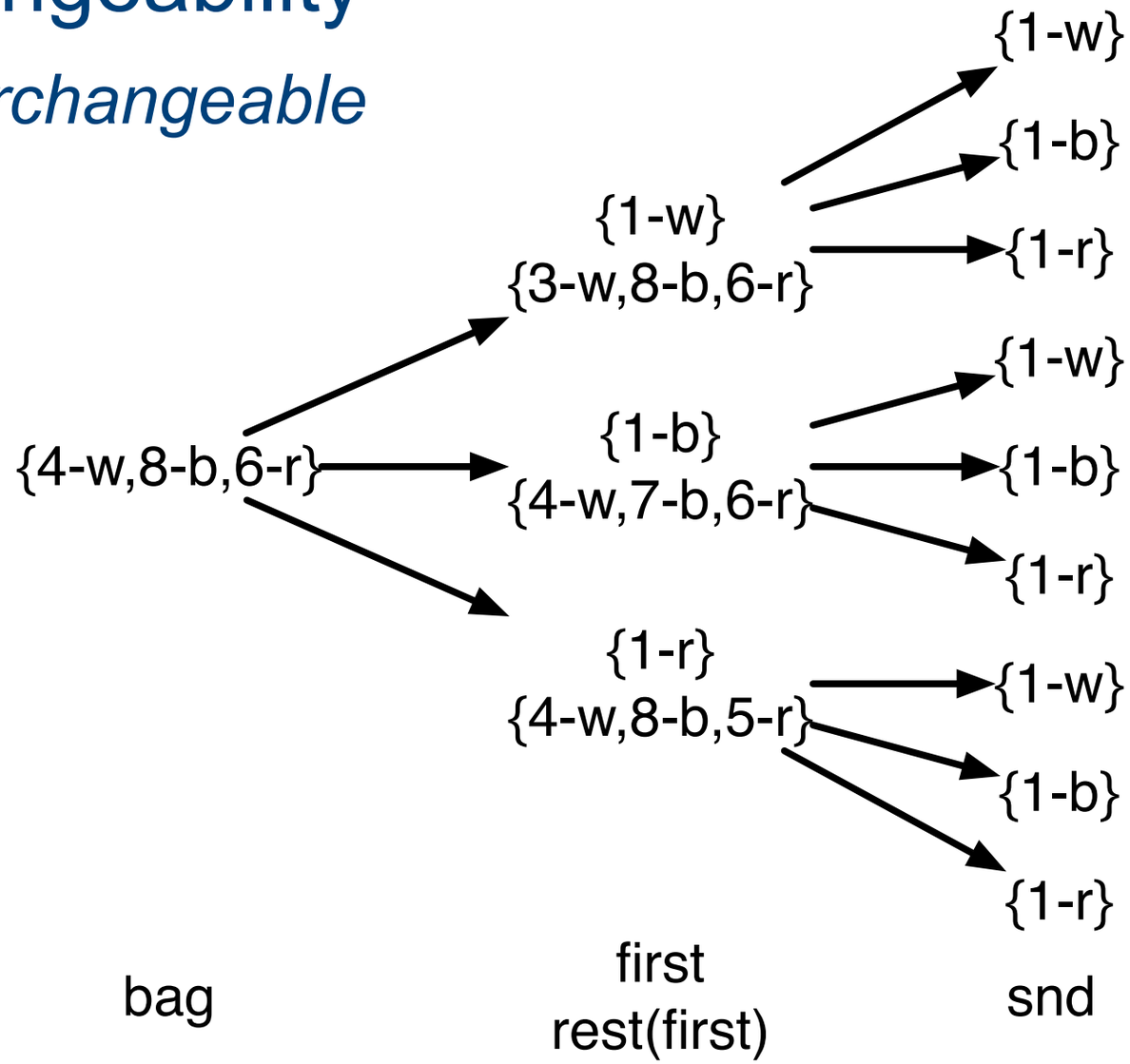
combinatorial explosion!

Solver

take into account exchangeability

all white balls are interchangeable

“lifted reasoning”



Solver

eliminate non-relevant information

A deck of cards consists of 4 aces, 4 twos, …,
4 queens and 4 kings.
If we take two cards at random,
what is the probability they are both queen?

A deck of cards consists of 52 cards of which 4 queens.
If we take two cards at random,
what is the probability they are both queen?

Solver

constraint-based reasoning on input

```
group(pop).

given(exactly(rel(94/100, pop, no_disease), pop, and(negative, no_disease))).
given(exactly(rel(96/100, pop, disease), pop, and(disease, positive))).
given(exactly(rel(2/100, pop), pop, disease)).

take(pop, person, 1).
observe(all(person, negative)).

probability(all(person, no_disease)).

property(test, [positive, negative]).
property(has, [no_disease, disease]).
```

#D = 0.02

#D + #H + #X = 1

#(H ∩ N) = 0.94 #H

#(D ∩ P) = 0.96 #D

→ #X = 0 and #H/#N=…

Solver

takes advantage of underlying solver technology

```
4/18::white(1); 8/18::blue(1); 6/18::red(1).

3/17::white(2); 8/17::blue(2); 6/17::red(2) :- white(1).
4/17::white(2); 7/17::blue(2); 6/17::red(2) :- blue(1).
4/17::white(2); 8/17::blue(2); 5/17::red(2) :- red(1).

evidence(red(1)).
query(white(2)).
```

```
6/18::red(1).

4/17::white(2) :- red(1).

evidence(red(1)).
query(white(2)).
```

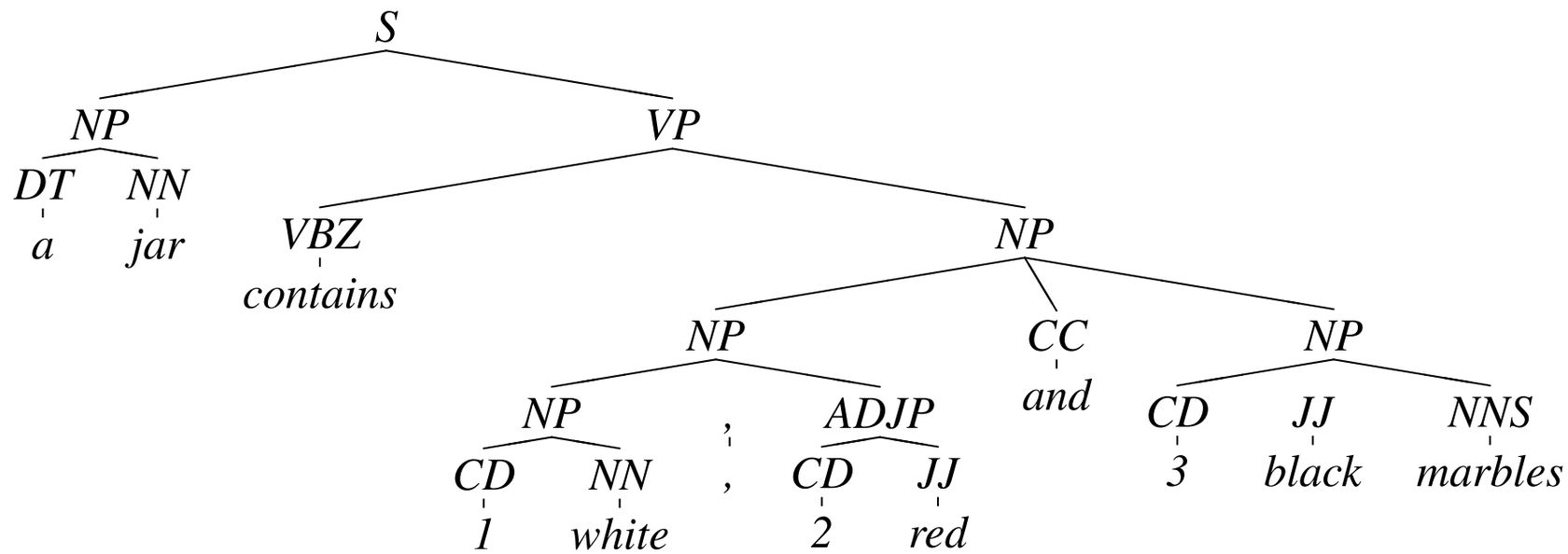
Challenges

1. Understand the question

2. Obtain background knowledge

3. Solve the question

NLP extraction



1. Find numbers
    2. Classify them
        3. Extract descriptive fragments
            4. Find question and post-process

! very naive approach !

Data

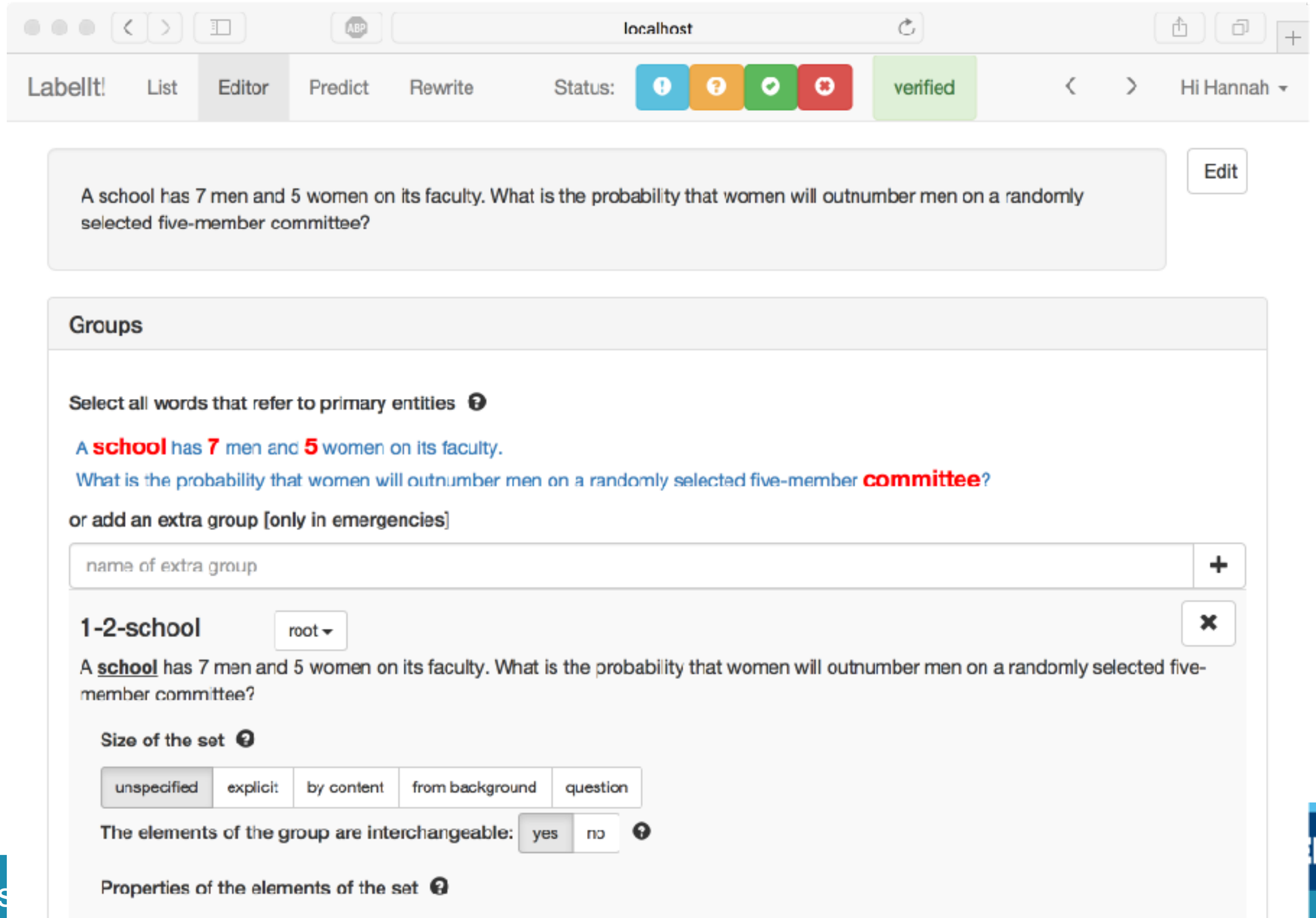
2376 questions on probability

structure of the problems labeled and solved by hand

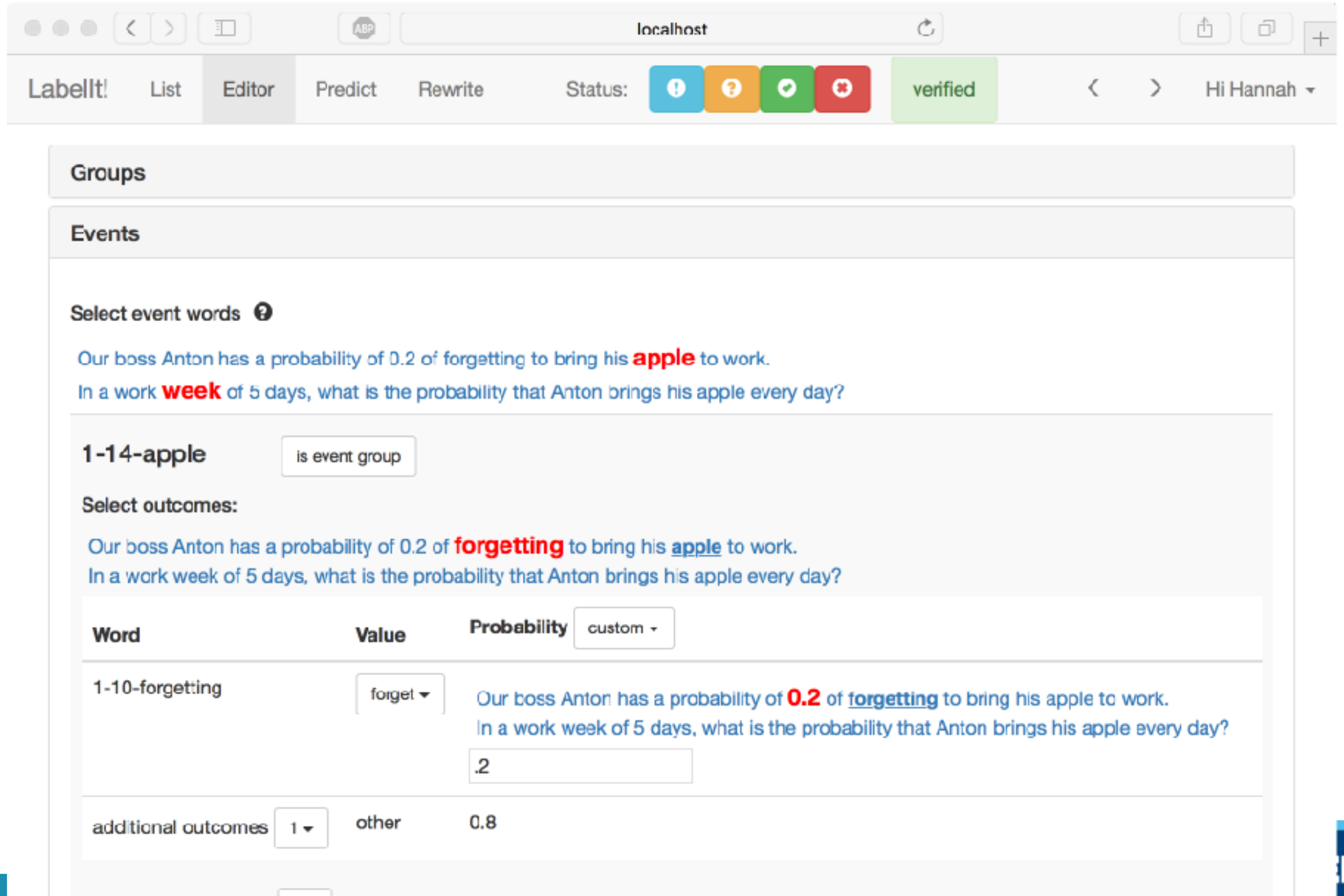
groups of objects

probabilistic events

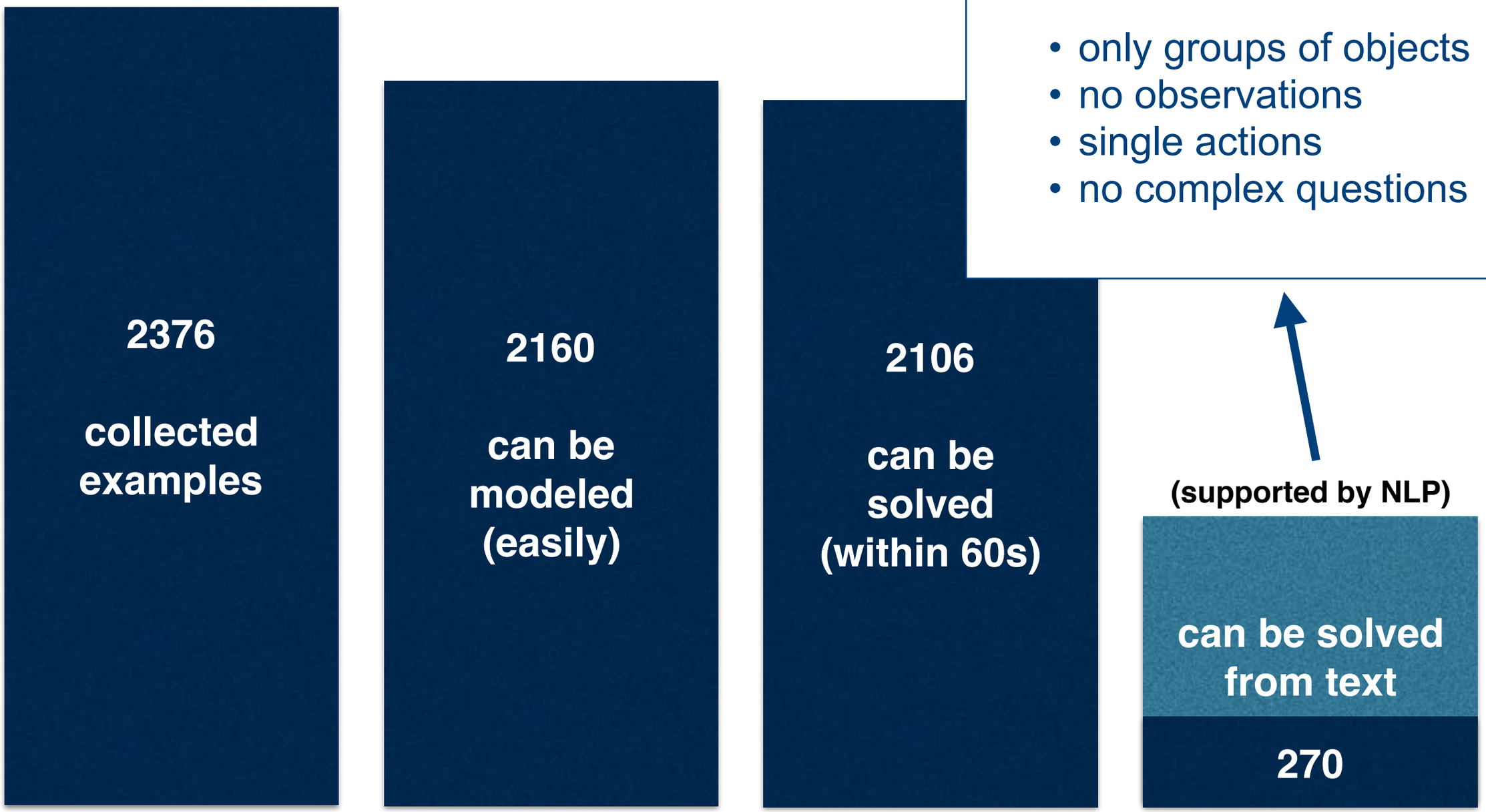
Data


localhost
LabelIt!
List
Editor
Predict
Rewrite
Status:
verified
Hi Hannah
A school has 7 men and 5 women on its faculty. What is the probability that women will outnumber men on a randomly selected five-member committee?
Edit
Groups
Select all words that refer to primary entities
A school has 7 men and 5 women on its faculty.
What is the probability that women will outnumber men on a randomly selected five-member committee?
or add an extra group [only in emergencies]
name of extra group
1-2-school
root
A school has 7 men and 5 women on its faculty. What is the probability that women will outnumber men on a randomly selected five-member committee?
Size of the set
unspecified
explicit
by content
from background
question
The elements of the group are interchangeable:
yes
no
Properties of the elements of the set
A school has 7 men and 5 women on its faculty.

Data


localhost
LabelIt!
List
Editor
Predict
Rewrite
Status:
verified
Hi Hannah
Groups
Events
Select event words
Our boss Anton has a probability of 0.2 of forgetting to bring his apple to work.
In a work week of 5 days, what is the probability that Anton brings his apple every day?
1-14-apple
is event group
Select outcomes:
Our boss Anton has a probability of 0.2 of forgetting to bring his apple to work.
In a work week of 5 days, what is the probability that Anton brings his apple every day?
Word
Value
Probability
custom
1-10-forgetting
forget
Our boss Anton has a probability of 0.2 of forgetting to bring his apple to work.
In a work week of 5 days, what is the probability that Anton brings his apple every day?
.2
additional outcomes
1
other
0.8
This event is repeated
5
times.
This event belongs to event group
2-4-week

Results



Limitations:

- only groups of objects
- no observations
- single actions
- no complex questions

Conclusion

1. a **dataset** of 2376 labeled questions

2. a **formal model** for representing probability questions
can represent 90.9% of collected questions

3. a **solver** for solving questions specified in the formal model
can solve 97.5% of modeled questions

Future work

4. an NLP component for **extracting a formal model from text**
can extract a correct model for 12.5% of questions

Thank you

KU LEUVEN

https://dtai.cs.kuleuven.be/problog/natural_language